

Introduction to PUMS: Public Use Microdata Sample

Chuck Purvis
Metropolitan Transportation
Commission, Oakland, California

Introduction & Applications

® Session #1: Introduction to PUMS

- What is PUMS?
- PUMS Geographic Areas (PUMAs)
- Top-Coding & Rounding Issues
- Software to Analyze PUMS
- Providing Bay Area PUMS Data to Users

® Session #2: PUMS Applications

- MTC Applications with 1980 & 1990 PUMS
- Possible Applications using PUMS 2000
- Ideas for County Planners & Transit Operators

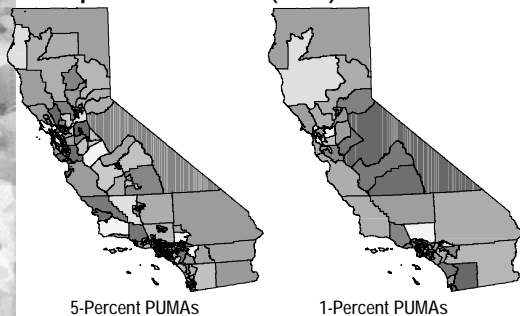
What is PUMS?

- Public Use Microdata Sample
 - Individual, disaggregate census "long form" data
 - Geo-coded ONLY at the PUMA level to protect confidentiality of respondent
 - One-percent PUMS (National File)
 - Five-percent PUMS (State File)
 - Evolving program, with 100,000+ population PUMAs starting in 1990
 - 1990 5% PUMS for Bay Area: Data for 108,000 households and 292,000 persons

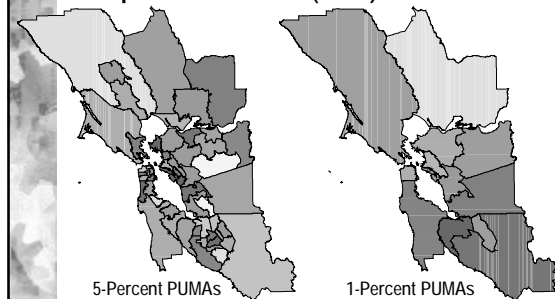
What are PUMAs?

- Public Use Microdata Areas
 - 5-percent sample PUMAs (100,000+ pop.)
 - 1-percent sample super-PUMAs (400,000+ pop.)
 - Defined by State Data Centers working with Regional Data Centers (MTC/ABAG) and others (UC Berkeley)
- California PUMAs:
 - 62 super-PUMAs; 233 PUMAs
- Bay Area PUMAs:
 - 12 super-PUMAs; 54 PUMAs

California PUMAs (5%) and Super-PUMAs (1%)



Bay Area PUMAs (5%) and Super-PUMAs (1%)



Top-Coding & Rounding (I)

- ® Income Top-Coding, use statewide mean for highest income groups
- ® Dollar Amounts (income, housing costs)

No Income	\$0
\$1 - \$7	\$4
\$8 - \$999	Round to nearest \$10
\$1,000 - \$49,999	Round to nearest \$100
\$50,000 +	Round to nearest \$1,000

Top-Coding & Rounding (II)

- ® Age, top-coded at 90, persons are assigned statewide mean for persons age 90+
- ® Race, Ethnicity, Ancestry
 - ® Nationwide minimum of 10,000 in group
- ® Departure time from home-to-work
 - ® 5-minute rounding for 5:00-10:59 AM
 - ® 10-minute rounding for 3:00-4:59; 11:00-midnight
 - ® 30-minute rounding for midnight – 4:59
- ® Travel times from home-to-work are rounded to nearest 5 minutes

PUMS 2000 Features

- ® Occupation based on Standard Occupational Classification (SOC)
- ® Industry based on the North American Industry Classification System (NAICS)
- ® Place of Work will be PUMA-of-Work
 - ® Super-PUMA to Super-PUMA commute flows in 1-percent PUMS
 - ® PUMA to PUMA commute flows in 5-percent PUMS
 - ® 1990 PUMS had PUMA-of-Residence to County-of-Work

Software to Analyze PUMS

SAS	www.sas.com	\$5000+
SPSS	www.spss.com	\$1150
Stata	www.stata.com	\$1200
Minitab	www.minitab.com	\$1200
S-Plus	www.insightful.com	
Systat	www.systat.com	\$1300
Statistica	www.statsoft.com	\$800
Shazam	shazam.econ.ubc.ca	\$400

List of Commercial & Free Statistical Software, at www.statistics.com
Also, check out www.stattransfer.com for conversion utility.

Where to Get PUMS?

- ® Integrated PUMS Project (Univ. Minnesota)
 - ® <http://www.ipums.org/>
 - ® <http://www.ipums.org/~census2000/>
- ® University of Virginia's GeoStat Program
 - ® <http://fisher.lib.virginia.edu/pums/>
- ® Census Bureau PUMS Program
 - ® <http://www.census.gov/population/www/cen2000/pums.html>
- ® MTC / ABAG / State Data Center

When to Get PUMS 2000?

- ® 1-percent National PUMS files expected March 2003
- ® 5-percent Statewide PUMS files expected later in 2003